DOCUMENT RESUME

ED 273 113 FL 015 928

AUTHOR Harrison, Andrew
TITLE Student-Centred Testing: Assessing Communication in Progress.
PUB DATE 84
NOTE 21p.; In: Willems, Gerard, Ed., Communicative Foreign Language Teaching and the Training of Foreign Language Teachers, 1984. p200-220. See FL 015 917.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Communicative Competence (Languages); Educational Strategies; *Interaction; *Language Tests; *Role Playing; Second Language Instruction; Student Centered Curriculum; *Test Construction

ABSTRACT

Assessment of communicative language learning means evaluating not only whether the language used is correct or incorrect, but also whether it is appropriate for the context. In structuring a communicative exchange for the purpose of testing, it must be determined that the problem used as the subject of an assessment interaction is suitable, and that it is at the appropriate level of language use, of interest to students, relevant to their language needs, and generalizable to other situations. The hypothetical problem should create an information gap and the incentive to bridge it. The linguistic, semi-linguistic, and non-linguistic means used to communicate in the exchange must be evaluated according to the alternatives available, the appropriateness of the language for the occasion, the degree to which error can be tolerated before it inhibits communication, the extent to which the essential elements of the exchange are present, and the degree to which it approaches native performance. The use of transcripts can be helpful in the analysis of spoken exchanges. Role-playing exchanges by fluent speakers can be used to establish the likely content and form of an interaction, for both evaluation and feedback to the student. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED273113

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gerard Willems

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

STUDENT-CENTRED TESTING: ASSESSING COMMUNICATION IN PROGRESS

Andrew Harrison

This article starts from a platitude: that teaching procedures and assessment procedures depend on each other. Tests should in theory look like a further application of a suitable learning method, not appear at the end of a learning process as elements in an alien system of value judgements. A coherent method of language teaching makes assumptions about a theory of learning, and these assumptions are reflected in its test procedures. The view that communication is the main aim of language learning assumes that working with examples of language used in a context, followed by progressively more complex applications, is a more valuable procedure than either description of a structural framework followed by ways of filling it out (the grammar/translation system), or the demonstration of a pattern followed by practice with vari-

FL 015 928

ations on it (the audio-lingual or audio-visual system). Tests resulting from the former approach ask students to process language according to the rules they have learnt, by translating to and from the foreign language, writing various kinds of essay, responding in an oral interview, reading aloud and writing from dictation. In each case the assessment system invites error and then penalises it when it occurs. The audio-lingual or -visual approaches lead to 'discrete point' tests which build up a picture of the student's skills like a mosaic from a large number of individual assessments, each testing a particular sound, structure or item of vocabulary. The basis for judging a student's success in this case is statistical probability established experimentally with trial versions of the tests. A third assessment procedure, which can fit both these theories of learning but is applied differently depending on which theory is adopted, is to offer the student a stretch of language and ask him to interpret the text in some way, or to put him in a situation and ask him to produce relevant language for these predetermined circumstances. Examples of the former kind of assessment are tests of listening and reading comprehension based on passages, cloze and dictation. Examples of the latter are tests of controlled writing (e.g. letters), oral narratives based on pictures series, and role plays.

These are brief descriptions of the three trends in testing labelled by Spolsky (1978) 'pre-scientific', 'psychometric-structuralist' and 'integrative-sociolinguistic', and characterized by Morrow (1979) as the Garden of Eden, the Vale of Tears and the Promised Land. The categories, whatever label is attached to them, are somewhat arbitrary because there are some elements which are common to all, depending on the point of view of the test writer, but they are helpful because they indicate the links which exist between teaching strategies and the corresponding principles for assessment. For the present discussion they provide a context for asking the question: What are the principles for assessment which follow on from a communication approach to language teaching and learning? But before considering any answers, an important distinction must be made between the association of teaching technique

with testing technique and the relationship between a testing technique and the reason for using it. One particular testing format may be appropriate for a number of different assessments: for example the dictation can be useful in any of the three approaches, depending on how it is constructed (words, sentences or passages); administered (teacher reading aloud or tape of authentic speech); and marked (by error, by sense groups or by single words restored to gaps in a written version of the text).

A communicative approach to teaching and learning assumes that the most important principle for deciding what should be learnt is the use of language in realistic contexts. When the tests mentioned above are looked at from a point of view of the student taking them, it becomes clear that very few of them require him to act as a language user in a real-life situation. The only possibilities are role play, letter writing and, for those whose livelihood may depend on it, translation from the foreign language. But real life is an impossible criterion for testing because a test is always to some extent a simulation. A real test of communicative performance might involve something like sending the student from London to Edinburgh with ten pence. If he reaches Edinburgh within a given time limit he has passed the test because he must have communicated successfully to get there: if he disappears somewhere in the Midlands he has failed. This kind of operational version of assessment cannot work however, even if more realistically set up. There are far too many extraneous factors which have nothing to do with communication (like luck, the weather, and above all the character of the student), and what communication there is, occurs far too randomly. For an assessment to produce useful results(without for the moment defining what 'useful' means) there must be enough control over the content for the necessary comparisons to be made. The control does not need to be in the form of a detailed specification, though it can be some kind of assessment, but should set the limits for what is acceptable in terms of both appropriateness and correctness for the purpose which is represented by the test.

The problem of how this control is to be achieved may be tackled by asking a defining question: 'How can we assess what is happening when language is used purposefully?'. The four important elements in this sentence are assess, happen, language and purpose.

Taking these in the reverse order - the logical one - the purpose of language use is communication, which means that there are at least two people involved. This is obviously true of a spoken exchange, but it is also true of communications through writing and reading. An entry in a polician's diary is not just a personal note of what colleagues said in Cabinet on a particular day, and this becomes obvious as soon as the diarist is out of office and the work is published as a justification of his past attitudes and actions. Language is written with a reader in mind, and is spoken for a listener. The reader or listener in turn responds to the meaning of the language he is offered. He reads for information or pleasure, listens in order to respond - and the results of this response may be very varied, ranging from a black eye arising from the reaction to an insult, to a note on a message pad as the outcome of a telephone call, as well as the more predictable development of a conversation as it progresses. It is not just a question of integrating the 'four skills' - of admitting the impossibility of separating reading from writing and listening from speaking - but a question of accepting that language is integrated with the circumstances in which it is used. The element of language in the defining question is therefore the language which is appropriate to these circumstances - where it is being used, by whom, with what aim in view and even with what success. What is happening is a sequence of language which arises from, and is modified by, the context in which it occurs.

An example of a sequence in this sense might be as follows: I am going to buy some bread. I do not need to read 'Baker & Confectioner' on the shop front because I know where our local bread shop is, but even if I didn't, I can see the bread in the window. I go in and say 'Hello' to a friend who happens to be in the shop, following this up with an enquiry about his wife's health, since I am surprised to find him buying bread instead of being at work. He explains that she is in bed with

flu: I commiserate. I reach the counter. 'Good morning', I say to the shop assistant, consulting my shopping list. 'Can I have a large Farmhouse and a small brown please?' "Sorry, no small browns left', she says. 'I'd better have a large one than', I reply, annotating the shopping list to explain the discrepancy to my wife later... And so the sequence continues. This scene is commonplace to the point of cliché, the kind of role play which will bear little or no repetition in class, but it contains a great deal of material for discussion.

The first point is that analysing the sequence from the point of view of one of the participants ('I'), separating out the communication skills involved, the semantic content and the result of interaction between these skills and content, shows that there is great variety even within this sort space. This is clear from Figure 1. Secondly, all four of the traditional language skills (Listen, Speak, Read, Write) are interwoven in the sequence. In addition, it appears that

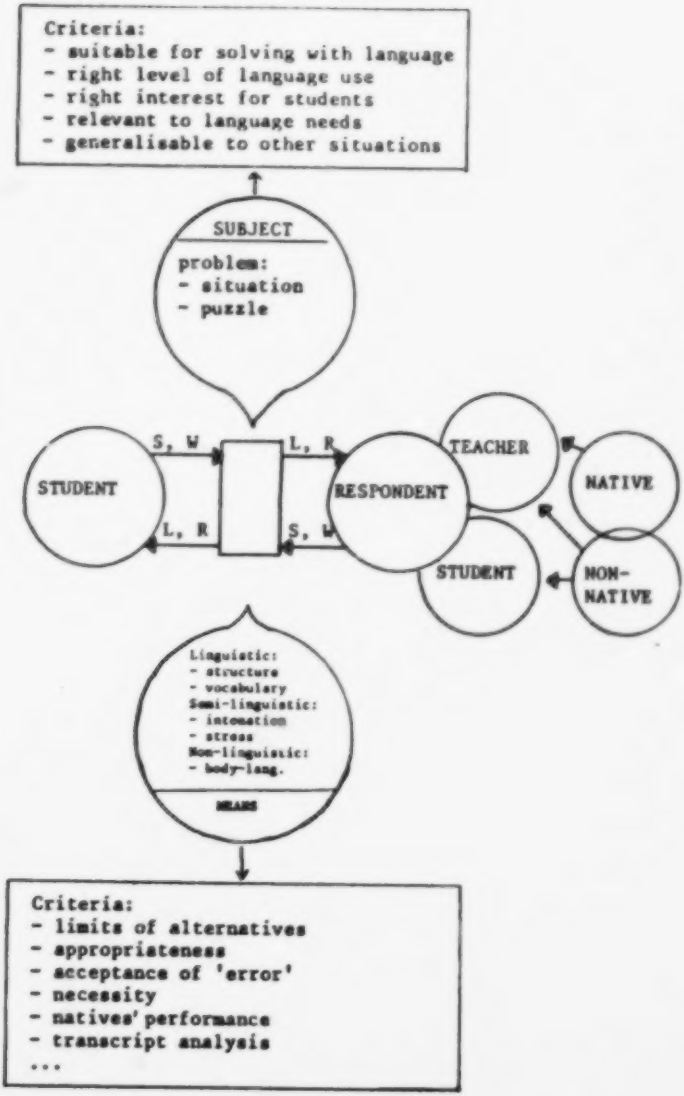
Figure 1 The bread shop sequence

skill needed by 'I'	content (semantic)	result (receptive)	intention (productive)
Read	Baker & Confectioner	(information)	
visual	bread in window	confirmation	
visual	friend	surprise,pleasure	
Speak	hello		greeting
Speak	how is wife?		enquiry
Listen	flu	information	
Speak	sorry		sympathy
visual	assistant	recognition	
Speak	good morning		greeting
Read	shopping list	reminder	
Speak	Farmhouse, small brown		request, purpose
Listen	sorry, no small		apology,information
visual	large one	information suggestion	
Speak	large one		request, purpose
Write	large		information

considerably more visual communication occurs in real life than is reflected in role plays and tests as they are commonly organised. I am surprised and pleased to see my friend, but would not start speaking to him unless I saw him; I see by general movement of people in the shop and by the fact that the assistant is looking at me that it is my turn to be served; my alternative request for a large loaf is triggered by seeing one on the shelf. Thirdly, the verbal exchanges involve three people, not just two, and the role relationship differs between each possible pair within the three. I greet my friend and the shop assistant in different ways, and it is clear (though not explicitly stated in the analysis in Figure 1) that my whole conversation with him is more informal than my transaction with her. (To clarify the relationship in the third pairing, I should add that my friend finds the bread shop girl quite attractive.) Fourthly, the transaction goes wrong at one stage (there are no small brown loaves), which means that a standard response is not enough to meet the situation. Communication involves coping with a sequence of events which does not necessarily follow the patterns of a drill practice: the student must be capable of applying the language he has learnt to a new situation by responding appropriately to it. Finally, the analysis of skills and results/intentions derived from the language used in the bread shop sequence are a more helpful guideline for setting up exchanges of this kind as role plays (either for teaching or for testing) than the language actually used on this particular occasion.

These five principles are helpful in considering how to assess communication. A sixth principle will sum up the discussion so far: what needs to be assessed is not just whether the language used by the candidate is right or wrong in itself, but whether it is appropriate, that is, right or wrong for the context.

Figure 2. A communicative test.



All these arguments are the same as those in favour of communicative (or functional/notional) language learning. How can they be applied to assessment? This is the first element in the defining question discussed earlier. The starting point is the information gap which is to be bridged by communication. This is the blank rectangle in the middle of Figure 2. This diagram relates particularly to the immediate, live exchange between speakers, but in principle the same approach can be applied to the time-lapsed communication which occurs with writing and reading. In general, what the diagram shows is that decisions have to be made about the subject of an exchange and the means of doing it, and also the criteria on which these decisions need to be made. The student is on one side of the information gap, the respondent on the other. (There is a difficulty in finding the right title for this other person because he may initiate rather than merely respond, and cannot be called an interlocutor if he does not actually speak, as when the student is instructing him in some action which he is to perform.) The respondent may be a teacher representing a native (or non-native) speaker in the real-life communication represented by the test; or he may be another student, who by definition can represent only a non-native speaker; or he may be one of a group who interact in turn with the student and each other. The student Speaks or Writes into the information gap and Listens and Reads out of it. The respondent matches these skills with the opposite ones: L and R for the student's S and W, and vice versa. Other skills which are not mentioned in the diagram but could be called into play are, on the productive side (with S and W), action and drawing - responses which are not linguistic but are evidence of understanding; and on the receptive side (with L and R), visual material - a prompt for production. In addition, the student's Reading may be of material which is not supplied by the respondent but is available as data (often in the form of realia) for the discussion across the gap.

Above the gap is the subject of the exchange - what the participants are talking about - and beneath, the means with which they do it - what they are communicating with. The subject for a test of communication

must set some kind of problem which creates an information gap and at the same time provides the incentive to bridge it. The problem is at its most realistic as a situation, which says in effect to the student: 'Suppose you were....' (in the bread shop/at a party/planning a holiday/furnishing your flat, and so on). But the possibilities for setting up realistic situations of this kind are limited, particularly at the earlier stages of language learning, so an additional strategy is proposed, labelled 'puzzle' in the diagram. This says to the student 'Here is a task... '(follow these instructions/solve this enigma/find out what is required/make sense of these elements, and so on). The puzzle is not realistic, in the sense that it is likely to coincide with the eventual purposes for which the student is learning the language, but it gives him a legitimate reason for using the language he has learnt for useful purposes such as discussion, agreement, or persuasion. It is, in short, functional.

What kinds of problem can be set for assessing communication? Five citeria are listed at the top of the diagram. First, it is obvious that the problem has to be one which requires the use of spoken or written language to solve it. A taxi can be hailed with no more than a wave of an umbrella; a game of bridge can be won by playing particular cards for the information of one's partner or, at worst, by aiming a kick under the table. Language as learnt in the classroom is superfluous in these circumstances. Secondly, the language level required must be suitable for the intended students. This is not merely a question of how much the candidate has learnt - whether he is a beginner or an advanced student - but what he will be required to do with the language which the test offers him and demands of him. It is possible, and if the real life situation is being represented, desirable, for him to be given authentic material to work with, but requiring less of him in response at the lower levels than at the higher levels, both in quantity and in quality of understanding. The third criterion, that the material used should be of the right interest for the students, again seems obvious, but as well as the subject matter (the words of a pop song as against a text on macro-economics) there are considerations of

the level of attack ('tourism' means something very different to an adult learner and a school pupil, though it is ostensibly the same topic) and the intrinsic interest of the problem which is set (is there some psychological pressure on the student to find the solution, or is the problem something he would not care to solve even in his native language?). The 'needs' argument is that if students are learning the language for a purpose, the language that they learn should be relevant for this purpose: in the same way, it is rather pointless to present a candidate with a role play about planning a holiday trip to the mountains if he is not going to visit the country where the language is spoken but is interested in reading technical literature. Finally the assessment must represent something which has wider application than the occasion of the test. It could be argued that a student who could cope with the bread shop sequence (in a more formalised, test version) could also cope with a similar transaction in a department store, even though the context is different, provided he knew the necessary vocabulary. It could hardly be maintained, however, that success in that sequence would forecast adequate performance at a company board meeting where a take-over deal is to be discussed, even though the transaction is essentially the same, a buying and selling procedure. This is perhaps an absurd example, but the generalisability of assessments is too easily taken for granted: it is in practice one of the more problematical aspects of validity and sampling.

The means for fulfilling the demands made by a test are both linguistic and nonlinguistic, and need no further elaboration at this point. The balloon in the diagram signals merely the fact of their existence. The selection of means for any particular assessment is guided by the criteria listed in the box under the 'means' balloon. The first criterion is how to set limits to the infinite possibilities for language performance. As suggested earlier, it is not feasible for the assessor to say simply: 'Speak and I will judge', because then it is impossible for him to make any but fortuitous comparisons between one performance and the next. And comparison is the essence of assessment. Different comparisons are required for different purposes: the student may want

to compare his present performance with that of a few weeks ago, or with the performance of other students in the same class, or more generally with students at the same level on a national basis; or to compare his language capability with the tasks it has been called upon to perform. Another crucial comparison is between the judgements of different assessors, either those made by different assessors on the same occasion or those made at different times by the same assessor. The only way in which these aspects of reliability can be tackled is to set some limitations to the language which is permissible in the test. One of the traditional methods of providing this control is to make the student show understanding of meaning in one language by expressing it in another: that is, translation. Another is to offer him a closed set of responses, as in a multiple-choice test. In a communicative test, the control is provided by the context.

It has already been suggested that one of the most important aspects of the language used in an exchange is its appropriateness for the occasion on which it is used. It is here that the problem of judging the effectiveness of a communication has to be faced. Appropriateness does not necessarily entail correct grammatical form, but it is concerned with whether the language used fits the roles of the speaker and his hearer (stranger/stranger, friend/friend), whether it is consistent within the established linguistic context of the exchange (a question of collocation, the right words in the right places and in the right order) and also whether it is suitable for the physical context. These conditions are the basis for deciding on the limits of the alternatives to be allowed - the essential ground rules for a marking scheme.

Another difficult question then has to be answered: how far can error be tolerated? The answer for a communicative test should be: until it stops communication. The point at which this happens depends on the other party to the exchange. For example, if a writing test consists of a letter written by the candidate in answer to an advertisement about accommodation, the marking scheme should presumably be based on the likely reaction of the recipient. If he is in business to promote the letting of hotel rooms and other accommodation to foreign tou-

rists, he is not concerned about whether 'Dear Sir' is followed by 'Yours faithfully' or 'Yours sincerely' though the former is the convention required by the rules of business correspondence. Still less is he interested in the spelling of 'sincerely'. He may have enough experience in the business to know what a German speaker means by 'I only need one room and a badroom'. But if dates are omitted or expressed ambiguously he will have to write back and ask for further information. At this point it is clear that communication has broken down. With this kind of blunt judgement, discrimination between student's performances is impossible to achieve except at the lowest levels. The student who walks into the hi-fi shop, points at the cassettes he wants and says 'Two' has communicated perfectly well for his purposes. He has not made any grammatical error and his vocabulary, as far as it goes, is perfect. There are arguments for rejecting the concept of 'error' in language learning altogether in favour of 'idiosyncratic dialects' since there is no evidence that the learner is failing to follow a known rule, only that the rules he has learnt so far do not apply in the sentence he is constructing ('Every sentence is to be regarded as idiosyncratic until shown to be otherwise', Corder 1974: 166). This is another route to appropriateness, since the judgement which a communicative assessment has to make about the quality of the student's performance relates not just to the formulation of acceptable sentences but also to the transmission of an appropriate response. It should be possible therefore to draw up a marking scheme which sets up categories of acceptance/rejection in terms of communication and appropriateness , with the latter including functional and social criteria which may well override grammatical and semantic 'correctness' as they are usually conceived of in language tests.

It has been suggested above that controls have to be set for the language which is to be produced by the student. This implies a limitation of the amount of material to be assessed. Conversely, however, it also implies a minimum which is to be produced. The criterion of communication on its own could lead to the acceptance of anything at all which seemed even basically communicative (for example 'Me, beer') in any

Germanic-language-speaking community), but this kind of exchange falls short of the limits which have to be set to allow the necessary comparisons to be made. The limits cannot be the perfect performance expected by a grammar/translation or structuralist approach, but must relate to the linguistic/communicative task in hand. A possible way forward is to look for the essentials in an act of communication and set those up as the basis for the marking. If the marking scheme can guide the assessor to make judgements on those elements which are necessary for the communication to take place, the language without which the exchange cannot reach a conclusion, everything else which happens can be left out of the assessment. This is an essential condition for a positive marking scheme. For example, in the bread shop sequence, it is communicatively successful and structurally correct, but socially inappropriate, for me to greet my friend with 'Where's your wife?' Although it is a well-formed enquiry from the purely linguistic point of view it is not an expression of friendly greeting, and in addition it would be interpreted as an insult because I know my friend's wife's name perfectly well but choose not to use it. For an assessment which was concerned with basic strategies for making contact with others there could be a marking scheme which laid down rules for communication first and appropriateness second. A student who produced a sentence like this in the context given would be credited with achieving communication but not appropriateness. But if an essential component of the assessment in the bread shop sequence were the function 'greeting', the student could be given no credit for this utterance as communication, and appropriateness would not be considered.

In setting up assessments on the basis of the essentials needed to communicate, the immediate problem is how to establish what these essentials are. The answer, almost inevitably, is the performance of native speakers. But proposing the native speaker as the ideal performer is too restrictive for this purpose, except perhaps for assessments at the highest levels: all that is needed are competent performances by fluent speakers - no more in fact than what is required of a teacher. If the task set in the test, whether it is a situation or a puzzle, is

solved on a trial basis by a number of fluent speakers, it should be possible to analyse what they say and decide on some essentials which can form the basis for assessment.

The analysis of performances presents further complications. Firstly since every fluent performance will be different, how many are needed before the essentials can be established? And secondly, what sort of analysis is required? The first of these questions has to be answered empirically, because the amount of variation in the performances which solve a communication problem depend on the nature of the problem itself. The range of possible conversational gambits which could fulfil the function 'getting to know someone at a social gathering' is very wide indeed, but necessarily includes greeting, giving and receiving information, stating likes and dislikes and so on. Describing how to perform some physical operation will probably require only instruction, confirmation and correction. In the first case, a larger number of fluent performances will be needed to establish what is necessary language content than in the second. The analysis of these performances does not need to be as thorough as the systems proposed for the analysis of discourse (Coulthard 1977), but it does need some theoretical basis so that a series of consistent decisions can be made about the requirements of a marking system.

The intention of the analysis is limited, not exhaustive: it is to find out which factors are to be assessed, as these have been established in the trial exchanges. The use of transcripts is helpful because the act of listening with understanding is normally a search for meaning, not for form or function, and it is only when the actual words used are available on a longer term basis than speech that an adequate analysis can be carried out. The result looked for depends on the purpose of the assessment, and this takes the discussion right back to the beginning, where the specification for the test asks and then tries to answer questions about purposes, as the basis for decisions about content.

In the present discussion the assumption has been that the purpose of a

communicative test is largely to assess proficiency - the ability to apply the language that has been learnt to situations which are subsequently met. The content should presumably therefore be expressed in functions. But the analysis could equally well look for what makes a function work in a given situation - the 'enabling skills' described by Morrow (1979). These are the contents of the 'means' balloon in Figure 2. Alternatively the analysis could be concerned with some combination of functions, structures and vocabulary. This raises several issues which cannot be dicussed in detail here, but may be briefly signalled. First, a function can often be fulfilled by so many different structures that it may be difficult if not impossible to lay down any 'necessary structures'. Second, the vocabulary used in a particular exchange may well not be generalisable, since the analysis is looking for applied lexis rather than the more traditional basis of assessment such as 'width' of vocabulary, which implies knowledge of rarer words and facility with reformulations. Third, the process of isolating the necessary means in a transcript (or a series of them) may seem like a return to the discrete-point approach, which is inappropriate for communicative tests because individual items without context do not add up to valid assessment of language for useful purpose. But in this case it is the context which makes the point acceptable or not. In any case, it will be possible to accept alternatives (which is not allowed with the traditional type of discrete-point test) since variations are possible within a given situation.

The above discussion has centred on the elements in a live exchange, since it is here that the greatest difficulty exists in trying to identify essentials. The same principles however apply to the time-lapsed exchange through the written word. The important point is that the relationship between the parties to an exchange is what governs the content of that exchange. The next stage is to attempt to put these suggestions into practice.

The principle of establishing the content of an exchange by noting what happens when fluent speakers perform it has been tries out experimentally with small grous of participants (The use of 'perform' is deli-

berate, since the experimental situation is even more unreal than an actual test.). The performance itself need not take long, since the length of an exchange depends on the nature of the task to be done, and some kinds of task may be completed in a few minutes. The aim of the exercise is to see how far the 'necessary' language can be extracted from repeated performances. Two members of the group play the parts in the exchange, as explained to them on individual cards (not seen by anyone else in the group), and the others act as observers, each one looking for a particular aspect of the exchange and noting it as it occurs, for discussion later. For example, one listens for those structures used by the two role players which seem essential for communication to take place; another listens for essential vocabulary; another notes functions expressed; another watches for body language used; and so on, depending on the number of participants. The amount of material in each of these categories and its importance relative to the others varies according to the demands of different tasks.

An example of a role play used for this exercise is a mock telephone conversation called 'An evening out?'. The task is based on an instruction card and two role cards, as reproduced in Figure 3. The intention is to provoke an exchange in which the players have different preferences and different aims in view, but are each open to persuasion by the other. The question mark in the title indicates that the players are not expected to come to any preconceived conclusion: indeed, they might in the end agree not to go out together this Saturday. The discussion itself is more important than the result arrived at. This may at first sight seem oddly at variance with the principles of setting a task to be achieved with language, and with the criterion-referenced assessment which logically follows: neither the task nor the assessment can be set up without specifying in advance what successful performance consists of. But the purpose of 'An evening out?' as a test would be to give the students a chance to show whether they can discuss , persuade, argue, put a point of view, and so on, in the context of a personal, friend/friend exchange. The vocabulary they use (which is largely given in the role cards) is less important in this instance

BEST COPY AVAILABLE

Figure 3. An exchange.

4. An evening out?

Take one card each. Sit back to back and use the information on the cards as the basis for a telephone conversation

An evening out?

No 1

You are Chris.

It is Friday evening and you have been looking at the adverts for local cinemas (see over) You telephone your friend Pat to ask him/her to go to see a film with you tomorrow evening.

You have had a busy week at work and want to relax and cheer yourself up. You have heard that "Elephant Man" is good, but you don't know much about it. You have already seen "Close Encounters"

What are you going to do on Saturday evening?

An evening out?

No 2

You are Pat.

Your telephone number is 997 4455.

It is Friday evening and you aren't doing anything on Saturday evening, but there is a quite good play on TV.

You think "Elephant Man" is a depressing film: it's about an ugly man and his unhappy life. You would like to see "Close Encounters" because you enjoyed it before and someone has told you that the new version is even better.

As a non-smoker, you hate the smell of cigarettes, and you know that the ABC cinema does not have a non-smoking section.

What are you going to do on Saturday evening?

than the functions which the players need to express and the structures which they find from their own language resources to achieve this expression.

For example, the beginning of the exchange is bound to include establishing the identities of the speakers, an explanation of why the telephone call is being made, information from the newspaper cutting, an offer or suggestion from No 1, and so on. The structures will include question forms, various tenses ('I've been looking at the paper'; 'Would you like to go?'; 'I've seen it'), negatives, comparatives and so on. The content of the exchange can be listed in the same way as the bread shop sequence (Figure 1), and this can be used as the basis for a marking scheme. In this way, the assessment is made positively, since priority is given to communications achieved rather than to the exact use of particular forms. The mark scheme might, for example, specify as one of the points for assessment the expression of past reference, and the student could achieve credit for this whether he said 'I've seen it' or 'I saw it last week , or avoided the past form of the verb altogether by saying 'I went to see it last week'. This avoidance strategy raises an important principle: the assessor should give it credit as expressing the required concept (or notion), not penalise it as an evasion. If the marking is done on the basis of credit for 'communication', plus credit for 'correctness' when communication has been achieved, these three examples would all be awared both. Grammatical errors would result in the loss (or rather non-award) of the 'correctness' credit. The performance of each student thus builds up into a series of achievements, which are interlocked in a sequence but are not necessarily dependent on each other, since the progress of the exchange is supported by both participants - it is in their interests that it should continue.

This kind of assessment system has considerable flexibility. It can be applied to different contents by adjusting the extent of the individual points which are to be assessed, from particular required structures to complete gambits such as an instruction or explanation. It can be

used for different purposes of testing, most obviously for proficiency tests, in which what has been learnt is applied to new situations, but also for diagnostic tests, which show whether the student has learnt what he has just been taught. It also has possibilities for self-assessment by providing for comparisons to be made with the student's previous efforts.

The approach to assessment through setting up communicative exchanges is not new in itself, nor does it offer possible solutions for assessment problems in general. The variety of situations in which assessments have to be made is so wide that a range of techniques is needed to cope with them. But where the interest lies in how well the student has learnt to apply language to communicative purposes at whatever level of sophistication he has reached, the procedures suggested above can give helpful guidelines in specifying and setting up the test situations. Perhaps most important of all, the student needs this kind of positive feedback about his progress to show him that the elements of language which he is learning are building up into a practical repertoire of communicative skills.

Bibliography

Corder, S.P., "Idiosyncratic dialects and error analysis", in Richards J.C.(ed) Error Analysis, Longman, London, 1974.

Coulthard, M., An introduction to discourse analysis, Longman, London, 1977.

Morrow, K.,"Communicative language testing: revolution or evolution?", in Brumfit, C.J. & Johnson, K. (eds), The communicative approach to language teaching, OUP, Oxford, 1979.

Spolsky, B. (Ed.), Approaches to language testing, Advances in language testing series: 2, Center for Applied Linguistics, Arlington, 1978.

Editors' note:

1. Readers might find it helpful to refer also to Schlesinger, J. & Carroll, M., (1980) Testing Communicative Competence, Oxford, Pergamon.

2. Andrew Harrison's above article was very much written with the EFL/ESL Secondary schoolclass in mind. In adult EFL/ESL teaching, but, to a certain extent, this also holds good for the school situation, it would be advisable in student-centred testing to make a distinction between external and internal assessment. In any communicative approach, the emphasis is placed, by definition, on the expression of the learners' intentions, needs and attitudes: subjective criteria for evaluation are therefore an essential part of learning and performing. If the reader is interested in this aspect of self-assessment in communicative foreign language testing, we refer him to:

Dickinson, L.,(1980) Report of the workshops on self-directed learning held under the auspices of the British Association for Applied Linguistics, Edinburgh, Scottish Centre for Overseas Education.

Holec, H.,(1980) Autonomy and foreign language learning. Oxford, Pergamon.

Oskarsson, M.,(1971) Self-assessment in foreign language learning. Oxford, Pergamon.